

## ON GEOMETRY OF NUMBERS

By HERMANN WEYL.

[Received 27 November, 1939.—Read 15 December, 1939.]

1. Minkowski stated his fundamental theorem concerning convex solids in relationship to a lattice as follows. Let the lattice in the  $n$ -dimensional space of vectors  $\mathfrak{x} = (x_1, \dots, x_n)$  consist of all vectors  $\mathfrak{a}$  with integral components, and let  $f(\mathfrak{x})$  be any "gauge function", i.e. any continuous function in vector space enjoying the following properties:

$$f(\mathfrak{x}) > 0 \text{ except for the origin } \mathfrak{x} = \mathfrak{o} = (0, \dots, 0);$$

$$f(t\mathfrak{x}) = |t| \cdot f(\mathfrak{x}) \text{ for any real factor } t;$$

$$f(\mathfrak{x} + \mathfrak{x}') \leq f(\mathfrak{x}) + f(\mathfrak{x}').$$

Then  $f(\mathfrak{x}) < 1$  defines an (open) convex solid  $\mathfrak{K}$  with  $\mathfrak{o}$  as its centre.  $V$  being the volume of  $\mathfrak{K}$  and  $2M$  the minimum of  $f(\mathfrak{a})$  for all lattice vectors  $\mathfrak{a} \neq \mathfrak{o}$ , Minkowski's universal inequality is

$$(1.1) \quad M^n V \leq 1.$$

If  $\mathfrak{S}$  is any set in vector space, we can place it round a given point  $\mathfrak{a}$  of the  $n$ -dimensional affine point space; the ensuing point set  $\mathfrak{S}_{\mathfrak{a}}$  consists of all points  $\mathfrak{x}$  such that the vector  $\mathfrak{x} = \overrightarrow{\mathfrak{a}\mathfrak{x}}$  is in  $\mathfrak{S}$ . Minkowski's argument is based on the simple remark that the convex solid

$$\mathfrak{K}(q): f(\mathfrak{x}) < q,$$

homothetic with  $\mathfrak{K}$ , when placed round each of the lattice points  $\mathfrak{a}$ , causes no overlapping so long as  $q \leq M$ , and hence up to that limit its volume  $q^n V$  stays less than or equal to 1. With  $q$  growing beyond  $M$ , the solids of the lattice begin to overlap. Studying the portion of the space covered by

them, Minkowski arrives at the sharper inequality†

$$(1.2) \quad M_1 \dots M_n V \leq 1,$$

where the numbers  $M_k$  may be described thus: A bubble of shape  $\mathfrak{K}$  while being blown up describes the series of figures  $\mathfrak{K}(q)$  with increasing  $q$ ;  $q = 2M = 2M_1$  is the first moment when a lattice point  $\mathfrak{b}_1 \neq \mathfrak{o}$  enters into the bubble,  $2M_k$  that stage at which for the first time a lattice point  $\mathfrak{b}_k$  enters off the  $(k-1)$ -dimensional linear manifold  $[\mathfrak{b}_1, \dots, \mathfrak{b}_{k-1}]$  spanned by  $\mathfrak{b}_1, \dots, \mathfrak{b}_{k-1}$ . Clearly  $M_1 \leq M_2 \leq \dots \leq M_n$ . In other words,  $\alpha = \mathfrak{b}_k$  is that one among the lattice vectors  $\alpha$  outside the manifold  $[\mathfrak{b}_1, \dots, \mathfrak{b}_{k-1}]$  for which  $f(\alpha)$  assumes the least possible value  $2M_k$ . So long as  $q \leq 2M_k$  the body  $\mathfrak{K}(q)$  contains less than  $k$  linearly independent lattice vectors, while for  $q > 2M_k$  their number is at least  $k$ .

The vectors

$$\pm \mathfrak{b}_1/2M_1, \quad \pm \mathfrak{b}_2/2M_2, \quad \dots, \quad \pm \mathfrak{b}_n/2M_n,$$

and hence their convex closure, are contained in  $\mathfrak{K}$ . If  $P = |\mathfrak{b}_1 \dots \mathfrak{b}_n|$  is the volume of the parallelotope spanned by the lattice vectors  $\mathfrak{b}_1, \dots, \mathfrak{b}_n$ , the volume of that closure is

$$\frac{2^n}{n!} \frac{P}{2M_1 \dots 2M_n} = \frac{P}{n!} \frac{1}{M_1 \dots M_n}.$$

Therefore

$$\frac{P}{n!} \leq M_1 \dots M_n V \leq 1,$$

and thus the positive integer  $P \leq n!$ .

This note consists of two parts, the first stating Minkowski's problem in terms of characteristic functions  $\phi$  which are capable of the two "truth values" 0 and 1 only, and then generalizing it to probability functions that move in the range  $0 \leq \phi \leq 1$ . Overlapping is not excluded, and is taken care of by using Boolean rather than ordinary sums. The substance of the argument remains unaltered, yet the essential points come out more clearly.

If  $P$  were equal to 1, the inequality (1.2) would tell an important fact about reduction under arithmetic equivalence, namely about normalization of the basis of our lattice in terms of a given convex solid  $\mathfrak{K}$ . As it stands, it points in a slightly different direction. However, by an almost trivial twist, one can make it tally with the problem of reduction. This is done in the second part, which concludes with some remarks about the special case of the reduction of quadratic forms.

---

† *Geometrie der Zahlen* (Leipzig, 1896, 1910), 211–218.

[Added January, 1940. Prof. L. J. Mordell, to whom I sent a copy of the MS., kindly pointed out to me two papers which I had overlooked; one by H. Davenport, *Quart. J. of Math.*, 10 (1939), 119–121, bears closely on Part I by containing another proof for the inequality (1.2), while the other by K. Mahler, *Quart. J. of Math.*, 9 (1938), 259–262, actually anticipates the main result of Part II, namely Theorem V. I am still surprised that the discovery of this fact in 1938 by Dr. Mahler, and of the simple remark on which it is based, had to wait for forty-two years after Minkowski established his inequality (1.2).]

## I.

2. A set in the affine space of points  $x = (x_1, \dots, x_n)$  may be described by its characteristic function  $\phi(x)$  which is equal to 1 inside the set and to 0 outside the set. The characteristic function  $\phi$  for the union of two sets with the characteristic functions  $\phi_1, \phi_2$  is their "Boolean" sum

$$\phi = \phi_1 + \phi_2 - \phi_1 \phi_2,$$

which reduces to the ordinary sum only if the sets do not overlap. Any number  $a$  in the interval  $0 \leq a \leq 1$  is said to be a *probability*. If  $a$  and  $b$  are probabilities of two statistically independent events  $A$  and  $B$ , then their Boolean sum

$$c = a \vee b = a + b - ab = 1 - (1-a)(1-b),$$

which again satisfies the inequality  $0 \leq c \leq 1$ , is the probability of the event " $A$  or  $B$ ". The Boolean sum is commutative and associative.  $a' \leq a$  implies  $(a' \vee b) \leq (a \vee b)$ , as the expression  $a \vee b = a(1-b) + b$  at once shows; and consequently

$$(2.1) \quad a' \leq a, \quad b' \leq b \quad \text{imply} \quad (a' \vee b') \leq (a \vee b).$$

The Boolean sum of a finite number of probabilities  $a_1, \dots, a_h$  is

$$\begin{aligned} \sum_i a_i &= a_1 \vee a_2 \vee \dots \vee a_h = 1 - \prod_{i=1}^h (1-a_i) \\ &= \sum_i a_i - \sum_{(i,k)} a_i a_k + \sum_{(i,k,l)} a_i a_k a_l - \dots \\ &= \frac{1}{1!} \sum_i a_i - \frac{1}{2!} \sum_{i,k} a_i a_k + \frac{1}{3!} \sum_{i,k,l} a_i a_k a_l - \dots \end{aligned}$$

Summation extends to all throws of 1 or 2 or 3 or ... different indices from the set 1, ...,  $h$ , in the second line without regard to their order, in the

third line with regard to their order. Brackets as in  $(i, k, l)$  indicate the first kind of summation.

A function  $\phi(x)$  in our space which satisfies the condition  $0 \leq \phi \leq 1$  is called a probability function.

Let  $\phi(x)$  now be an (integrable) probability function vanishing outside a finite region  $\mathfrak{H}$  of the  $x$ -space which may be fixed as a parallelotope:

$$A_i < x_i < B_i \quad (i = 1, \dots, n).$$

The lattice translation  $\mathfrak{a} = (a_1, \dots, a_n)$  with integral components  $a_i$ ,

$$x_i' = x_i + a_i,$$

carries  $\phi(x)$  into a function

$$\phi(x, \mathfrak{a}) = \phi(x_1 - a_1, \dots, x_n - a_n).$$

In these circumstances, the Boolean sum

$$\psi(x) = \mathbf{S}_{\mathfrak{a}} \phi(x, \mathfrak{a})$$

extending to all lattice translations  $\mathfrak{a}$  has a meaning, since for each  $x$  only a limited number of its terms do not disappear.  $\psi(x)$  is a probability function with period 1 in each of the variables  $x_1, \dots, x_n$ . We form the integral

$$J[\phi] = \int_E \psi(x) dx \quad [dx = dx_1 dx_2 \dots dx_n]$$

extending over the fundamental mesh of the lattice

$$E: \quad 0 < x_1 < 1, \dots, 0 < x_n < 1.$$

Obviously  $0 \leq \psi(x) \leq 1$  entails

$$(2.2) \quad 0 \leq J[\phi] \leq 1.$$

Moreover, for two probability functions  $\phi_2$  and  $\phi_1 \leq \phi_2$  we have  $\psi_1 \leq \psi_2$ , because of the law of monotony (2.1) for Boolean sums, and hence

$$(2.3) \quad J[\phi_1] \leq J[\phi_2].$$

The explicit expression for  $\psi$  is

$$\begin{aligned} \psi(x) &= 1 - \prod_{\mathfrak{a}} \{1 - \phi(x, \mathfrak{a})\} \\ &= \frac{1}{1!} \sum_{\mathfrak{a}} \phi(x, \mathfrak{a}) - \frac{1}{2!} \sum_{\mathfrak{a}, \mathfrak{b}} \phi(x, \mathfrak{a}) \phi(x, \mathfrak{b}) + \frac{1}{3!} \sum_{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}} \phi(x, \mathfrak{a}) \phi(x, \mathfrak{b}) \phi(x, \mathfrak{c}) - \dots \\ &\quad [\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \dots \text{ lattice vectors}]. \end{aligned}$$

In substituting  $a+b$ ,  $a+c$ , ... for  $b$ ,  $c$ , ... we see that  $\psi(x)$  may be written as an ordinary sum

$$\psi(x) = \sum_a \phi^*(x, a),$$

where

$$\phi^*(x) = \phi(x) \left\{ \frac{1}{1!} - \frac{1}{2!} \sum_b \phi(x, b) + \frac{1}{3!} \sum_{b,c} \phi(x, b) \phi(x, c) - \dots \right\}$$

[ $b, c, \dots$ , lattice vectors  $\neq 0$ ].

This trick works because the lattice translations form a *group*. When we return to "combinations" without regard to order, the second factor on the right in the last formula is

$$(2.4) \quad \phi^{**}(x) = 1 - \frac{1}{2} \sum_{(a)} \phi(x, a) + \frac{1}{8} \sum_{(a,b)} \phi(x, a) \phi(x, b) - \dots$$

[ $a, b, \dots$  lattice vectors  $\neq 0$ ].

Take a real parameter  $t$  ranging from 0 to 1, and introduce the product

$$\prod_{a \neq 0} \{1 - t\phi(x, a)\} = 1 - t \sum_{(a)} \phi(x, a) + t^2 \sum_{(a,b)} \phi(x, a) \phi(x, b) - \dots$$

Its integration from  $t = 0$  to 1 leads to (2.4). Hence our final definitions are:

$$(2.5) \quad \phi^*(x) = \phi(x) \cdot \phi^{**}(x), \quad \phi^{**}(x) = \int_0^1 \prod_{a \neq 0} \{1 - t\phi(x, a)\} dt;$$

$$(2.6) \quad J[\phi] = \int_{\infty} \phi^*(x) dx.$$

$\infty$  indicates that the integral extends over the whole  $x$ -space. Notice that  $0 < \phi^{**}(x) \leq 1$  and  $\phi_1^{**} \geq \phi_2^{**}$  if  $\phi_1 \leq \phi_2$ . (2.6) proves  $J[\phi]$  to be invariant under translation, namely

$$(2.7) \quad J[\phi'] = J[\phi] \quad \text{for} \quad \phi'(x) = \phi(x - x^0).$$

If  $0 \leq \phi_1 \leq \phi$  and  $\phi - \phi_1 \leq \epsilon$ , we find that

$$\phi^*(x) = \phi(x) \cdot \phi^{**}(x) \leq \phi(x) \cdot \phi_1^{**}(x)$$

and thus, from (2.6), derive the result

$$(2.8) \quad J[\phi] - J[\phi_1] \leq \int_{\mathfrak{H}} \{\phi(x) - \phi_1(x)\} \phi_1^{**}(x) dx \leq H\epsilon,$$

where  $H$  is the volume of  $\mathfrak{H}$ . We summarize:

**THEOREM I.** *The integral (2.6) is less than or equal to 1 for any probability function  $\phi$  which vanishes outside a finite region.  $J[\phi]$  is*

*invariant under translation, and*

$$J[\phi_1] \leq J[\phi_2] \quad \text{if} \quad \phi_1(x) \leq \phi_2(x).$$

Specilization of the theorem for the characteristic function  $\phi$  of an (open) set  $\mathfrak{S}$  of definite Jordan volume  $V$  yields the simple Minkowski-Blichfeldt-Scherrer theorem: If  $\mathfrak{S}_0$  has no point in common with the sets  $\mathfrak{S}_a$  placed round the other lattice points  $a \neq 0$ , or, what is the same thing: if  $\mathfrak{S}_0$  never contains both end points  $x$  and  $x+a$  of a lattice vector  $a \neq 0$ , then its volume is less than or equal to 1. I consider Theorem I to be the natural general form of this principle, and our proof is nothing but a neat analytical form of the simplest argument by which that principle has been proved†.

[Our affine  $x$ -space could be replaced by any differentiable manifold  $U$  of points  $x$  which carries an infinitesimal volume measure, and the group of lattice translations by any group of differentiable volume-preserving one-to-one transformations  $x \rightarrow x^s$  of  $U$  which

(1) is discrete in the strictest sense, so that the set of points  $x^s$ , equivalent to any given point  $x$ , never has a condensation point,

(2) has a compact fundamental domain; *i.e.*, the manifold, arising from  $U$  by identification of equivalent points, is supposed to be compact.]

3. With Siegel‡ we introduce the Fourier coefficients of the periodic function  $\psi(x)$ :

$$J_l = \int_E \psi(x) \cdot e(l_1 x_1 + \dots + l_n x_n) dx$$

corresponding to the various lattice-vectors  $l = (l_1, \dots, l_n)$  in the dual space, with  $e(x)$  as an abbreviation for  $e^{2\pi i x}$ , and formulate the Parseval equation

$$\int_E (\psi(x))^2 dx = \sum_l |J_l|^2.$$

The zero coefficient  $J_0$  is our  $J = J[\phi]$ . Making use of the equation

$$\int_E \psi(x) \cdot \omega(x) dx = \int_\infty \phi^*(x) \cdot \omega(x) dx$$

† G. Hajós, *Act. Litt. Sci., Szeged*, 6 (1934), 224–225. H. F. Blichfeldt, *Trans. American Math. Soc.*, 15 (1914), 227–235. W. Scherrer, *Math. Annalen*, 86 (1922), 99–107.

‡ *Acta Math.*, 65 (1935), 307–323.

which holds for any periodic function  $\omega(x)$ , we get

$$(3.1) \quad J_1 = \int_{\infty} \phi^*(x) \cdot e(l_1 x_1 + \dots + l_n x_n) dx$$

and

$$\int_E (\psi(x))^2 dx = \int_{\infty} \phi^*(x) \cdot \psi(x) dx.$$

We find that

$$\begin{aligned} \phi^* \psi &= \phi^* \left\{ 1 - \prod_{\alpha} (1 - \phi(x, \alpha)) \right\} \\ &= \phi^* - \phi^* (1 - \phi) \cdot \prod_{\alpha \neq 0} \{1 - \phi(x, \alpha)\} \end{aligned}$$

and thus

$$(3.2) \quad \begin{aligned} \phi^* \psi &= \phi^* - \phi(1 - \phi) \phi_* \phi^{**} \quad \text{with} \\ \phi_*(x) &= \prod_{\alpha \neq 0} \{1 - \phi(x, \alpha)\}. \end{aligned}$$

**THEOREM II.** *If  $\phi_*$  and  $\phi^{**}$  are defined by (3.2), (2.4) and  $J_1$  denotes the integral (3.1), then the following equation holds for  $J = J[\phi]$ :*

$$J(1 - J) = \int_{\infty} \phi(1 - \phi) \cdot \phi_* \phi^{**} dx + \sum_{\alpha \neq 0} |J_1|^2.$$

It shows very clearly how  $0 \leq \phi \leq 1$  leads to  $0 \leq J \leq 1$ .

[This argument works on any differentiable manifold  $U$  on which we have a transitive group  $\Gamma$  of differentiable one-to-one transformations, the lattice consisting of a discrete subgroup of  $\Gamma$  whose fundamental domain is compact. The harmonics replacing  $e(l_1 x_1 + \dots + l_n x_n)$  are obtained by a slight modification of the process described by the author, *Annals of Math.*, 35 (1934), 486–499.]

4. We now introduce an assumption of *convexity* by supposing  $\{\phi(x) \geq r\}_x$ , i.e. the locus  $\mathfrak{F}_r$  of the points  $x$  where  $\phi(x) \geq r$ , to be a convex set for any  $r > 0$ . The relations  $\phi(x) \geq r$  and  $\phi(x') \geq r$  then imply

$$(4.1) \quad \phi(tx + (1-t)x') \geq r$$

for any number  $t$  between 0 and 1. A function  $\phi$  having this property in our  $n$ -dimensional space preserves it on any linear subspace. The “terrace”  $\{r \leq \phi(x) < r'\}_x$  between the levels  $r$  and  $r' > r$  has a Jordan volume because

it is the difference between the two convex bodies  $\mathfrak{F}_r$  and  $\mathfrak{F}_r$ . Let

$$(4.2) \quad 0 = r_0 < r_1 < r_2 < \dots < r_h = 1$$

be any division of the interval  $0 \leq r \leq 1$  into subintervals of lengths  $r_{i+1} - r_i \leq \epsilon$ , and let  $\phi_\Delta(x)$ ,  $\phi^\Delta(x)$  be the corresponding step functions which approximate  $\phi$  from below and above with an error less than or equal to  $\epsilon$ . They vanish outside  $\mathfrak{H}$  and are defined inside  $\mathfrak{H}$  by

$$\phi_\Delta(x) = r_i \text{ and } \phi^\Delta(x) = r_{i+1} \text{ if } r_i \leq \phi(x) < r_{i+1}.$$

The Riemann integrals of  $\phi_\Delta$  and  $\phi^\Delta$  clearly differ by not more than  $H\epsilon$ ; this proves the Riemann integrability of  $\phi$ . Moreover, by (2.3) and (2.8),

$$J[\phi_\Delta] \leq J[\phi] \leq J[\phi^\Delta] \quad \text{and} \quad J[\phi^\Delta] - J[\phi_\Delta] \leq H\epsilon,$$

and these inequalities could also serve to define  $J[\phi]$  by mean of the Riemann integrals  $J[\phi_\Delta]$  and  $J[\phi^\Delta]$  of the approximate terraced functions.

We repeat our assumptions (A) about  $\phi$ :

- (i)  $0 \leq \phi(x) \leq 1$ ;
- (ii)  $\phi(x) = 0$  outside a finite region  $\mathfrak{H}$ ;
- (iii) the sets  $\mathfrak{F}_r: \{\phi(x) \geq r\}_x$  are convex.

Minkowski's second step deals with the process of dilation as defined by

$$(4.3) \quad \phi_q(x_1, \dots, x_n) = \phi\left(\frac{x_1}{q}, \dots, \frac{x_n}{q}\right)$$

and, properly generalized, leads to

**THEOREM III.** *Under the assumptions (A),*

$$(4.4) \quad J[\phi_q] \geq J[\phi] \text{ for } q \geq 1.$$

*Proof.* Suppose that we have a point  $x^0$  where  $\phi(x)$  takes on its maximum value. If  $x$  is any point and  $\phi(x) = r$ , we then have  $\phi(x^0) \geq r$ , and, by (4.1),

$$(4.5) \quad \phi\left(\frac{1}{q}x + \left(1 - \frac{1}{q}\right)x^0\right) \geq r = \phi(x) \text{ for any } q \geq 1.$$

For the moment denote the function on the left-hand side by  $\phi_q'(x)$ . Formula (4.5) entails

$$J[\phi_q'] \geq J[\phi],$$

and thus, since  $\phi_q'$  arises from  $\phi_q$  by a translation,

$$J[\phi_q] = J[\phi_q'] \geq J[\phi].$$

Since we have neither assumed  $\phi(x)$  to be continuous nor the sets  $\mathfrak{F}_r$  to be closed, the existence of a maximum is doubtful. But the following slight modification will do. Let  $r_i$  in the finite sequence (4.2) be the highest level for which  $\mathfrak{F}_{r_i}$  is not empty. We cut off the top of the mountain at the altitude  $r_i$ , i.e. we form

$$\hat{\phi}(x) = \min(r_i, \phi(x))$$

and choose  $x^0$  as a point in  $\mathfrak{F}_{r_i}$ . Then we get

$$\phi_q'(x) \geq \hat{\phi}(x) \quad (q \geq 1)$$

instead of (4.5), and hence

$$J[\phi_q] \geq J[\hat{\phi}].$$

However, according to (2.8)

$$J[\hat{\phi}] \geq J[\phi] - H\epsilon,$$

and by driving  $\epsilon$  towards 0 the inequality (4.4) is re-established. (Instead of the decapitated mountain, we might have used the terraced mountain  $\phi_{\Delta}$ .)

Since

$$(\phi_q')_q = \phi_{q'q},$$

we may claim Theorem III to assert that

$$J[\phi_q] = J(q)$$

*is a monotonically increasing function of the parameter  $q$  of dilation.*

5. Our next move is to modify the procedure by taking into consideration only the lattice vectors

$$(5.1) \quad (a_1, \dots, a_k, 0, \dots, 0) \quad (a_i \text{ integers, } k \geq 1)$$

in the linear subspace

$$x_{k+1} = \dots = x_n = 0.$$

$\phi^* = \phi^{(k)}$  is now defined by the last equation (2.5) with the product ranging over the vectors  $\alpha \neq 0$  of this  $k$ -dimensional lattice, and

$$J^{(k)}[\phi] = \int_{\infty} \phi^{(k)}(x) dx.$$

The integration is then conveniently carried out in two steps, first with respect to  $x_1, \dots, x_k$  for any constant  $x_{k+1}, \dots, x_n$ , and secondly with respect to  $x_{k+1}, \dots, x_n$ . Thus we find that

$$\phi_1 \leq \phi_2 \text{ implies } J^{(k)}[\phi_1] \leq J^{(k)}[\phi_2],$$

while the inequality

$$J^{(k)}[\phi] \leq J^{(k)}[\phi_1] + H\epsilon,$$

holding under the condition  $0 \leq \phi - \phi_1 \leq \epsilon$ , follows by the same argument as (2.8), without a break in integration.

Like  $\phi$  itself, the function

$$\Phi(x_1 \dots x_k) = \phi(x_1 \dots x_k, x_{k+1}^0 \dots x_n^0)$$

of  $k$  variables satisfies the assumptions (A), and therefore

$$(5.2) \quad J[\Phi_q] \geq J[\Phi] \quad \text{for } q \geq 1.$$

Notice that  $\Phi_q(x_1 \dots x_k) = \bar{\phi}_q(x_1 \dots x_k, x_{k+1}^0 \dots x_n^0)$

with  $\bar{\phi}_q(x) = \phi\left(\frac{x_1}{q}, \dots, \frac{x_k}{q}, x_{k+1}, \dots, x_n\right).$

Thus, by integrating (5.2) with respect to  $x_{k+1}, \dots, x_n$ , we get

$$J^{(k)}[\bar{\phi}_q] \geq J^{(k)}[\phi].$$

The substitution

$$(5.3) \quad \frac{x_{k+1}}{q} = x'_{k+1}, \dots, \frac{x_n}{q} = x'_n$$

performed on the last  $n-k$  variables results in the equation

$$J^{(k)}[\phi_q] = q^{n-k} \cdot J^{(k)}[\bar{\phi}_q].$$

Hence  $J^{(k)}[\phi_q] \geq q^{n-k} \cdot J^{(k)}[\phi] \quad \text{for } q \geq 1.$

For our Riemann integrals the performance of the integration in two steps  $x_1 \dots x_k | x_{k+1} \dots x_n$  and of the substitution (5.3) offers no difficulty. But maybe it is preferable first to argue for the step function  $\phi_\Delta$ , thus avoiding the difficulty about the maximum of  $\phi$  mentioned above, and then to combine the ensuing inequality

$$J^{(k)}[\phi_q] \geq q^{n-k} \cdot J^{(k)}[\phi_\Delta]$$

with

$$J^{(k)}[\phi_\Delta] \geq J^{(k)}[\phi] - H\epsilon.$$

For  $k = 0$  we define  $\phi^{(0)} = \phi$  and therefore

$$J^{(0)}[\phi_q] = q^n \cdot \int_{\infty} \phi(x) dx.$$

Any  $k$ -dimensional discrete lattice  $\mathfrak{L}_k$  may serve here instead of (5.1), since the first  $k$  coordinate vectors  $e_1, \dots, e_k$  may be so chosen as to span the pre-assigned lattice (that is, so that the lattice consists of the vectors  $a_1 e_1 + \dots + a_k e_k$  with integral components  $a_i$ ).

**THEOREM IV.** *By means of a  $k$ -dimensional lattice  $\mathfrak{L}_k$  we define*

$$\phi^{(k)}(x) = \phi(x) \cdot \int_0^1 \prod_a \{1 - t\phi(x, a)\} dt,$$

*the product extending to all vectors  $a \neq 0$  of the lattice  $\mathfrak{L}_k$ , and*

$$J^{(k)}[\phi] = \int_{\infty} \phi^{(k)}(x) dx, \quad J_k(q) = J^{(k)}[\phi_q].$$

*Then  $J_k(q)/q^{n-k}$  is an increasing function of  $q$ . In particular,*

$$J_0(q)/q^n = \int_{\infty} \phi(x) dx.$$

The application to the situation considered by Minkowski is immediate.  $\phi$  is now the characteristic function of the convex solid  $\mathfrak{K}$ :  $f(x) < 1$ . Let  $\mathfrak{L}_k$  consist of the lattice vectors in  $[\mathfrak{b}_1, \dots, \mathfrak{b}_k]$ . So long as  $q \leq M_{k+1}$ , we have

$$\phi_q^{(k)} = \phi_q^*, \quad J_k(q) = J(q).$$

Indeed, the solids  $\mathfrak{K}(q)$  and  $\mathfrak{K}_a(q)$  do not overlap if  $a$  is a lattice point outside  $[\mathfrak{g}_1, \dots, \mathfrak{g}_k]$  and  $q \leq M_{k+1}$ . We thus find that

$$J(q) = q^n \cdot V \quad \text{for } q \leq M_1,$$

in particular

$$J(M_1) = M_1^n \cdot V.$$

Then

$$J(q) \geq \left(\frac{q}{M_1}\right)^{n-1} \cdot J(M_1) = M_1 q^{n-1} \cdot V \quad \text{for } M_1 \leq q \leq M_2,$$

and hence

$$J(M_2) \geq M_1 M_2^{n-1} \cdot V.$$

Continuing in the same manner, we get

$$J(q) \geq M_1 \dots M_{k-1} q^{n-k+1} \cdot V \quad \text{for } M_{k-1} \leq q \leq M_k$$

$$(k = 1, \dots, n)$$

and ultimately

$$J(q) \geq M_1 \dots M_n \cdot V \quad \text{for } q \geq M_n.$$

Since  $J(q)$  always remains less than or equal to 1, Minkowski's inequality (1.2) is proved.

## II.

6. Contrary to the procedure in I, yielding the vectors  $\mathfrak{b}_k$  and ratios of dilation  $2M_k$ , the *problem of reduction* requires the determination of vectors  $\mathfrak{e}_k$  and ratios  $N_k$  according to the following inductive rule:

Let  $\mathfrak{a}$  range over all lattice vectors outside the  $(k-1)$ -dimensional manifold  $[\mathfrak{e}_1 \dots \mathfrak{e}_{k-1}]$  such that every lattice vector in the  $k$ -dimensional manifold  $[\mathfrak{e}_1 \dots \mathfrak{e}_{k-1} \mathfrak{a}]$ ,

$$\mathfrak{x}_{k-1} + x\mathfrak{a} \quad \text{with } \mathfrak{x}_{k-1} \text{ in } [\mathfrak{e}_1 \dots \mathfrak{e}_{k-1}],$$

has an integral  $\mathfrak{a}$ -component  $x$ .  $\mathfrak{a} = \mathfrak{e}_k$  is a vector within this range  $\mathfrak{R}_k$  for which  $f(\mathfrak{a})$  takes on the least possible value  $N_k$ .

As the first step, we may take  $\mathfrak{e}_1 = \mathfrak{b}_1$ , so that  $N_1 = 2M_1$ . By induction we readily conclude that  $\mathfrak{e}_1, \dots, \mathfrak{e}_k$  span (*i.e.* form a basis of) the whole lattice in the manifold  $[\mathfrak{e}_1 \dots \mathfrak{e}_k]$ ; hence  $\mathfrak{e}_1, \dots, \mathfrak{e}_n$  span the whole  $n$ -dimensional lattice. Relative to this coordinate system, the lattice vectors

$$\mathfrak{x} = x_1 \mathfrak{e}_1 + \dots + x_n \mathfrak{e}_n = (x_1, \dots, x_n)$$

are again described by integral components  $x_i$ , and our reduction is equivalent to the following conditions:

$$(6.1) \quad f(x_1, x_2, \dots, x_n) \geq f(\delta_1^k, \dots, \delta_n^k) = N_k$$

which hold provided that  $x_1, x_2, \dots, x_n$  are integers and  $x_k, \dots, x_n$  are without common divisor. The  $\delta_i^k$  are the Kronecker  $\delta$ 's and thus  $\mathfrak{e}_k = (\delta_1^k, \dots, \delta_n^k)$ , the  $k$ -th unit vector.

This characterization (6.1) obviously entails the inequalities

$$N_1 \leq N_2 \leq \dots \leq N_n.$$

We have normalized the basis of our  $n$ -dimensional lattice relative to the convex solid by means of the  $n$  sets (6.1) of infinitely many inequalities, and this is what the problem of reduction demands. The question arises whether one can establish a universal upper bound for  $N_1 \dots N_n V$ . I now prove

**THEOREM V.** *If a reduced lattice basis  $\mathfrak{e}_k$  is so chosen as to satisfy the conditions (6.1), then*

$$(6.2) \quad N_1 \dots N_n V \leq \mu_n,$$

where

$$(6.3) \quad \mu_n = 2^n \cdot \left(\frac{3}{2}\right)^{\frac{1}{2}(n-1)(n-2)}.$$

*Proof.* One of the vectors  $\mathfrak{d}_1, \dots, \mathfrak{d}_k$ , say  $\mathfrak{d}$ , lies outside the linear manifold

$$E_{k-1} = [\mathfrak{e}_1 \dots \mathfrak{e}_{k-1}].$$

In

$$E_k^* = [\mathfrak{e}_1 \dots \mathfrak{e}_{k-1} \mathfrak{d}]$$

we can find a lattice vector  $\mathfrak{e}^*$  which, together with  $\mathfrak{e}_1, \dots, \mathfrak{e}_{k-1}$ , spans the whole lattice in  $E_k^*$ . Out of the finite number of lattice vectors of the form

$$(6.4) \quad x_1 \mathfrak{e}_1 + \dots + x_{k-1} \mathfrak{e}_{k-1} + y \mathfrak{d}$$

with  $-\frac{1}{2} < x_i \leq \frac{1}{2} \ (i = 1, \dots, k-1), \ 0 < y \leq 1$ ,

select one, say  $\mathfrak{e}^*$ , with the lowest  $y$ . Since  $\mathfrak{d}$  itself must be an integral combination of  $\mathfrak{e}_1, \dots, \mathfrak{e}_{k-1}, \mathfrak{e}^*$ , this  $y$  is the reciprocal  $1/e$  of a positive integer  $e$ . If  $e = 1$ ,  $\mathfrak{e}^*$  coincides with  $\mathfrak{d}$ . Our  $\mathfrak{e}^*$  is in  $\mathfrak{R}_k$  and thus in the competition for the lowest value  $N_k$  of  $f(\mathfrak{a})$ . But, since  $f(\mathfrak{d})$ , as one of the numbers

$$2M_1, \dots, 2M_k,$$

is certainly less than or equal to  $2M_k$ , we infer from these remarks and (6.4) that

$$f(\mathfrak{e}^*) \leq 2M_k \text{ if } e = 1,$$

$$f(\mathfrak{e}^*) \leq \frac{2M_k}{e} + \frac{1}{2}(N_1 + \dots + N_{k-1}) \leq M_k + \frac{1}{2}(N_1 + \dots + N_{k-1}) \text{ if } e \geq 2.$$

Hence  $N_k$  cannot exceed the larger of the two numbers

$$(6.5) \quad M_k + \frac{1}{2}(N_1 + \dots + N_{k-1}) \quad \text{and} \quad 2M_k.$$

This allows us to establish universal inequalities of the kind

$$(6.6) \quad N_i \leq 2\theta_i M_i.$$

Indeed (6.6) is true for  $i = 1$  with  $\theta_1 = 1$ . Once it holds good for

$$i = 1, \dots, k-1,$$

the bound (6.5) for  $N_k$  yields a similar inequality for  $i = k$  with  $\theta_k$  as the greater of the two numbers

$$1 \quad \text{and} \quad \frac{1}{2}(1 + \theta_1 + \dots + \theta_{k-1}).$$

This determines the factors  $\theta_k$  by recursion, and we readily verify that

$$\theta_1 = 1, \quad \theta_k = \left(\frac{3}{2}\right)^{k-2} \quad \text{for } k \geq 2.$$

Therefore (1.2) implies

$$N_1 \dots N_n V \leq 2^n \cdot \theta_1 \dots \theta_n = \mu_n.$$

It would be an interesting problem to ascertain the lowest constant  $\mu_n$  for which the relation (6.2) holds.

Fuller exploitation of the method yields the following

**GENERALIZED THEOREM V.** *Let  $q_1, \dots, q_n$  be given positive numbers greater than or equal to 1. If the inequality*

$$f(x_1, x_2, \dots, x_n) \geq \frac{1}{q_k} \cdot f(\delta_1^k, \dots, \delta_n^k)$$

*holds whenever  $x_1, \dots, x_n$  are integers and  $x_k, \dots, x_n$  have no common factor ( $k = 1, \dots, n$ ), then the values*

$$N_k = f(\delta_1^k, \dots, \delta_n^k)$$

*satisfy the relations*

$$q_k N_{k+1} \geq N_k \quad (k = 1, \dots, n-1)$$

*and*

$$N_1 \dots N_n V \leq 2q_1 \dots q_n (1+q_1)^{n-1} \bullet (1+\frac{1}{2}q_2)^{n-2} \dots (1+\frac{1}{2}q_{n-1})^1.$$

The proof is practically the same as before. The recursive equations for the  $\theta_k$  are now

$$\theta_1 = q_1 \quad \text{and} \quad \frac{2\theta_k}{q_k} = 1 + \theta_r + \dots + \theta_{k-1} \quad (k \geq 2)$$

$$\text{i.e.} \quad \frac{2\theta_2}{q_2} = 1 + q_1, \quad \frac{2\theta_k}{q_k} + \theta_k = \frac{2\theta_{k+1}}{q_{k+1}} \quad (k \geq 2),$$

with the solution

$$\theta_k = \frac{1}{2}q_k(1+q_1) \bullet (1+\frac{1}{2}q_2) \dots (1+\frac{1}{2}q_{k-1}) \quad (k \geq 2).$$

7. Here is another similar proposition:

**THEOREM VI.** *Put*

$$h_\nu = \nu^2 - \frac{1}{2}\nu - 1 \quad (\nu = 1, 2, \dots)$$

*and*

$$\rho_n = 2^{2\nu-n} \cdot (2\nu+2)(2\nu+3) \dots (\nu+n+1) \quad \text{for} \quad h_\nu \leq n \leq h_{\nu+1}.$$

*We can determine a lattice basis  $e_1^*, \dots, e_n^*$  such that the distances  $f(e_k^*) = M_k^*$  satisfy the inequality*

$$(7.1) \quad M_1^* \dots M_n^* V \leq \rho_n.$$

This constant  $\rho_n$  is much lower than the one in Theorem V. Its increase with  $n$  is best judged by the recurrence formulae

$$\rho_n = \frac{n+\nu+1}{2} \rho_{n-1} \quad \text{if } 1+h_\nu \leq n \leq h_{\nu+1},$$

and 
$$\rho_n = \frac{(n+\nu)(n+\nu+1)}{2(n+\nu)+1} \rho_{n-1} \quad \text{for odd } \nu \text{ and } n = h_\nu + \frac{1}{2}.$$

*Proof.* Out of the vector basis  $\mathfrak{d}_1, \dots, \mathfrak{d}_n$ , we construct a lattice basis  $e_1^*, \dots, e_n^*$  according to the recurrent equations

$$(7.2) \quad e_k^* = \frac{1}{e_k} \mathfrak{d}_k + a_{k,k-1} \mathfrak{d}_{k-1} + \dots + a_{k,1} \mathfrak{d}_1$$

with the conditions

$$-\frac{1}{2e_i} < a_{k,i} \leq \frac{1}{2e_i} \quad (i < k),$$

the  $e_k$  being positive integers. If  $e_k = 1$ , all the coefficients  $a_{k,k-1}, \dots, a_{k,1}$  vanish, and  $f(e_k^*) = 2M_k$ . However, if  $e_k \geq 2$ ,

$$M_k^* = f(e_k^*) \leq \frac{2}{e_k} M_k + \sum_{i=1}^{k-1} M_i / e_i \leq \eta_k M_k$$

with 
$$\eta_k = \frac{2}{e_k} + \left( \frac{1}{e_{k-1}} + \dots + \frac{1}{e_1} \right) \quad (e_k \geq 2).$$

To cover both cases we put  $\eta_k = 2$  for  $e_k = 1$ . From (1.2) we now get

$$M_1^* \dots M_n^* V \leq \eta_1 \dots \eta_n,$$

and it remains to discuss the product  $\eta_1 \dots \eta_n$ .

We put

$$\delta_k = 2 \text{ or } \delta_k = 1 \quad \text{according as } e_k = 1 \text{ or } e_k \geq 2;$$

$$\eta_k' = 2 \text{ if } \delta_k = 2, \text{ and } \eta_k' = \delta_k + \frac{1}{2}(\delta_{k-1} + \dots + \delta_1) \text{ if } \delta_k = 1.$$

Then 
$$\eta_k \leq \eta_k' \quad \text{and} \quad \eta_1 \dots \eta_n \leq \eta_1' \dots \eta_n'.$$

Among the  $2^n$  values of the product

$$(7.3) \quad \eta_1' \dots \eta_n'$$

resulting from the  $2^n$  possible sequences  $\delta_1, \dots, \delta_n$  of 1's and 2's, we have to ascertain the largest. An interchange of two consecutive figures 1 and 2 in this sequence whereby

$$\delta_k = 1, \delta_{k+1} = 2 \quad \text{is changed into} \quad \delta_k = 2, \delta_{k+1} = 1$$

(all other figures remaining unaltered) affects merely the two factors  $\eta_k' \eta_{k+1}'$  in (7.3), turning their values

$$[1 + \frac{1}{2}(\delta_{k-1} + \dots + \delta_1)] \cdot 2 \quad \text{into} \quad 2 \cdot [2 + \frac{1}{2}(\delta_{k-1} + \dots + \delta_1)],$$

and so increasing (7.3). For a given number  $\nu$  of 2's in the sequence, the maximum value is thus attained when they come first ( $\delta_1, \dots, \delta_\nu$ ), and then (7.3) is equal to

$$2^\nu \cdot \frac{2\nu+2}{2} \cdot \frac{2\nu+3}{2} \dots \frac{2\nu+1+\mu}{2} \quad (\mu = n - \nu)$$

i.e.  $2^{2\nu-n} \cdot (2\nu+2)(2\nu+3) \dots (\nu+1+n) = w_\nu.$

Let  $\nu$  pass through the values  $\nu = 0, 1, \dots, n$ . Since

$$w_\nu = \frac{4(\nu+1+n)}{2\nu(2\nu+1)} \cdot w_{\nu-1},$$

$w_\nu$  is on the upgrade ( $w_\nu > w_{\nu-1}$  or  $w_\nu \geq w_{\nu-1}$ ) so long as

$$\nu(2\nu+1) < 2(\nu+1+n) \quad \text{or} \quad \nu(2\nu+1) \leq 2(\nu+1+n),$$

and then goes down. Hence the maximum is attained for the last  $\nu$  for which  $\nu(\nu - \frac{1}{2})$  does not exceed  $n+1$ .

Here is a table for the lowest values of  $\rho_n$ :

$$\begin{array}{cccccccc} n = 1, & 2, & 3, & 4, & 5, & 6, & 7, & \dots \\ \rho_n = 2, & 4, & 12, & 42, & 168, & 756, & 3960, & \dots \end{array}$$

Higher bounds, for instance

$$\rho_n = (n+1)! \quad \text{or} \quad \rho_n = 2 \cdot n!,$$

can even more easily be obtained. Again we have the question of the best value of  $\rho_n$  in the inequality (7.1).

### 8. A positive-definite quadratic form

$$(8.1) \quad f(x) = \sum g_{ik} x_i x_k \quad (i, k = 1, \dots, n)$$

is said to be *reduced* if it satisfies the inequalities

$$(8.2) \quad f(x) \geq g_{kk} \quad \text{for every } x \text{ in } \mathfrak{R}_k \text{ and } k = 1, \dots, n,$$

where  $\mathfrak{R}_k$  is the set of all lattice points  $x = (x_1, \dots, x_n)$  for which  $x_k, \dots, x_n$  are without common divisor. While for every positive form of discriminant  $D = \det (g_{ik})$ ,

$$g_{11} \dots g_{nn} \geq D,$$

we have found that for a reduced form

$$(8.3) \quad \lambda_n g_{11} \dots g_{nn} \leq D \quad \text{with} \quad \lambda_n = (\omega_n/\mu_n)^2.$$

Here  $\omega_n$  denotes the volume of the  $n$ -dimensional unit sphere, most handily described by the recurrent relation

$$\omega_n = \frac{2\pi}{n} \cdot \omega_{n-2} \quad \text{together with} \quad \omega_0 = 1, \quad \omega_1 = 2.$$

Indeed, with  $\sqrt{f}$  as gauge function, the convex body  $\mathfrak{K}: f < 1$ , becomes an ellipsoid of volume  $\omega_n/\sqrt{D}$ , as is readily proved by linear transformation of  $f(x)$  into the normal form  $x_1^2 + \dots + x_n^2$ .

The linear substitutions  $S$  of the variables  $x_i$  with integral coefficients  $s_{ik}$  of determinant  $|s_{ik}| = \pm 1$  (or only those of determinant 1) form a group  $\{S\}$ . Two quadratic forms  $f$  and  $Sf$  which arise from each other by a substitution of the group are said to be *equivalent*. Every positive form is equivalent to a reduced form.

All quadratic forms  $f = \{g_{ik}\}$  with real coefficients  $g_{ik} = g_{ki}$  form an  $N = \frac{1}{2}n(n+1)$ -dimensional linear space  $R$ , the positive ones an open convex region  $G$  in  $R$ , the reduced forms a convex subset  $Z$  of  $G$  which is closed relative to  $G$ . Both  $G$  and  $Z$  are cones and may, therefore, be considered as sets in the  $(N-1)$ -dimensional sphere of rays in  $R$  issuing from the origin. A convex cone  $H$  is a set having these three properties:

- (i)  $f = 0$  is not in  $H$ ;
- (ii) if  $f$  is in  $H$ , so is  $tf$ , where  $t$  is any positive factor;
- (iii) with  $f$  and  $f'$  in  $H$ , the sum  $f+f'$  is in  $H$ .

We operate throughout in  $G$ ; all topological words referring to subcones of  $G$  are meant relative to  $G$  (not  $R$ ).

Each of the inequalities (8.2) for the variables  $g_{ij}$  is of the form

$$(8.4) \quad \sum_{i,j} a_{ij} g_{ij} \geq 0$$

with integral coefficients  $a_{ij}$ . We expressly exclude from the set  $\mathfrak{J}$  of inequalities (8.2) those for which all coefficients  $a_{ij}$  vanish, namely (8.2) with  $x = \pm(\delta_1^k, \dots, \delta_n^k)$ .

Each individual inequality (8.4) of  $\mathfrak{J}$  defines a half-space bounded by a plane

$$(8.5) \quad \sum_{i,j} a_{ij} g_{ij} = 0.$$

A substitution  $S$  carries the central cell  $Z$  into an equivalent cell  $Z_S$  which is also convex. The equivalent cells jointly cover the whole space  $G$ .  $Z_J$  coincides with  $Z$  for any of the  $2^n$  substitutions

$$J: x_1 \rightarrow \pm x_1, \dots, x_n \rightarrow \pm x_n,$$

and  $S$  and  $SJ$  have the same effect upon  $Z$ \*

From (8.3) Minkowski deduced two important theorems of finiteness:

I. *A finite number of the inequalities  $\exists$  suffices to define the cell  $Z$ .*

II. *There is only a finite number of substitutions  $S$  capable of carrying a point of  $Z$  into a point of  $Z$*  (worded more precisely, the demand on the element  $S$  of  $\{S\}$  which limits it to a finite set is to the effect that there should exist two points  $f, f'$  in  $Z$  such that  $f' = Sf$ ).

Minkowski has a straightforward algebraic proof of II, in § 7 of his paper "Diskontinuitätsbereich für arithmetische Äquivalenz"†. His proof of I is involved because he establishes I simultaneously with an inequality of the nature of (8.3) by a complicated induction. This entanglement was dissolved in a joint paper by L. Bieberbach and I. Schur‡ (oh tempi passati!). They improve on I by breaking it up into the two theorems:

I\*. *There is only a finite number of lattice points  $x$  in  $\mathfrak{R}_k$  for which there exists a reduced form  $f_0 = \{g_0^0\}$  satisfying the equation*

$$(8.6) \quad f_0(x) = g_{kk}^0.$$

I\*\*. *If we substitute for  $x$  in succession each of the lattice points assigned by Theorem I\*, then the corresponding inequalities*

$$(8.7) \quad f(x) \geq g_{kk}$$

*for  $k = 1, \dots, n$  completely close in the cell  $Z$  in  $G$ .*

While the proof of I\* again is a straightforward algebraic affair based on (8.3), I\*\* depends on the topological argument that a straight line joining a point inside to a point outside a convex region must cross the border. I explain the Bieberbach-Schur argument for I\*\* in this simplified form.

† *Journal für Math.*, 129 (1905), 220–274; *Gesammelte Abhandlungen* 2 (Leipzig, 1911), 53–100.

‡ *Sitzungsber. Berlin Akad.*, (1928), 519–523; (1929), 508.

I say that those points  $f$  (in  $G$ ) for which all the inequalities  $\mathfrak{z}$  hold with the sign  $>$  rather than  $\geq$  form the *core* of  $Z$ . Any diagonal form

$$g_1 x_1^2 + \dots + g_n x_n^2 \quad \text{with} \quad 0 < g_1 < g_2 < \dots < g_n$$

evidently belongs to the core of  $Z$ . So do the inner points of  $Z$ . I maintain that conversely every point of the core is an inner point of  $Z$ . Indeed if  $f$  is a positive form,  $U_\epsilon$  a sufficiently small neighbourhood of  $f$  and  $A$  any positive number, then the points  $x$  of the  $x$ -space such that  $f_\epsilon(x) \leq A$  holds for at least one  $f_\epsilon$  in  $U_\epsilon$  belong to a limited region. Therefore all lattice points  $x$  with a finite number of exceptions satisfy the inequality  $f_\epsilon(x) > A$  for every  $f_\epsilon$  in  $U_\epsilon$ . This follows readily from Jacobi's step transformation of  $f_\epsilon$  into a square sum. Hence we have a finite subset  $\mathfrak{z}^0$  of  $\mathfrak{z}$  such that every form  $f_\epsilon$  in  $U_\epsilon$  satisfies all inequalities of  $\mathfrak{z} - \mathfrak{z}^0$ . Provided that the inequalities  $\mathfrak{z}^0$  hold for  $f$  with the  $>$  sign, we may shrink  $U_\epsilon$  so as to have the inequalities  $\mathfrak{z}^0$  satisfied throughout  $U_\epsilon$ . This establishes I\*\* by the topological argument mentioned above.

Let  $f_0$  again be a point on the boundary of  $Z$  and let  $x$  range over the finite number of lattice points  $x$  in  $\mathfrak{R}_k$  which satisfy (8.6). Studying the convex wedge defined by the corresponding simultaneous inequalities (8.7), we find† that those inequalities  $\mathfrak{z}$  suffice which are satisfied with the  $=$  sign for  $N-1$  linearly independent points  $f$  in  $Z$ . They are at the same time indispensable, and the corresponding equations describe the  $(N-1)$ -dimensional planes which contain the faces of the pyramidal cell  $Z$  in  $G$ .

Another way of constructing these faces would be by determining the common boundaries of  $Z$  with the adjacent equivalent cells  $Z_S$ . But  $Z$  and  $Z_S$  have a point  $f_0$  in common only if both  $f_0$  and  $S^{-1}f_0$  are in  $Z$ . Hence the finiteness of the number of plane faces must also be a consequence of II, so that I ought to be deducible from II. Let us examine the situation more carefully.

Only the substitutions  $J$  are capable of transforming a point in the core of  $Z$  into a point of  $Z$ . Hence the two convex cells  $Z$  and  $Z_S$  have no inner point in common provided that  $S$  is not a  $J$ . The points which they have in common, if any, form a convex cone  $W$  in a linear submanifold of  $N-1$  or ... or 1 dimensions. If  $N-1$  is the correct number of dimensions we speak of a wall separating  $Z$  from  $Z_S$ . Choose  $f_0$  as an inner point of  $W$  in its  $(N-1)$ -dimensional plane. It satisfies one of the inequalities  $\mathfrak{z}$ , say (8.4), with the sign  $=$  while every point of  $W$  satisfies it with the  $\geq$  sign. Therefore the whole  $W$  clearly lies in the plane (8.5), and thus I is again established, *provided that we can show that every point  $f$  on the boundary of  $Z$*

† For an elementary treatment see H. Weyl, *Comm. Math. Helvet.*, 7 (1934-5), 290-306.

lies on a wall separating  $Z$  from an adjacent cell  $Z_S$ . We deduce this from the fact, important in itself, that the equivalent cells cluster only towards the boundary of  $G$ .

On closer examination Minkowski's proof of II reveals that there is not more than a finite number of  $S$  capable of carrying a point of  $Z$  into a point of  $G(\rho, \mu)$ . Here  $\rho$  is any number greater than 1 and  $\mu$  is any positive number. If  $D_k$  denotes the determinant of the form  $f(x_1 \dots x_k 0 \dots 0)$ , the subcone  $G(\rho, \mu)$  of  $G$  is defined by the following simultaneous inequalities:

$$\begin{aligned} D_k &\geq \frac{\lambda_k}{\rho} g_{11} \dots g_{kk}, \\ g_{kk} &\geq \frac{1}{\rho} g_{k-1, k-1} \quad [g_{00} = 0], \\ f(x_1, \dots, x_{k-1}, 1, 0, \dots, 0) &\geq g_{kk} - \mu g_{11} \end{aligned} \quad (k = 1, \dots, n)$$

for any integers  $x_1, \dots, x_{k-1}$ .

All points of  $Z$  are *inner* points of  $G(\rho, \mu)$ . This is more than sufficient to settle our question. Choose a definite  $\rho > 1$  and  $\mu > 0$ . Let  $f$  be a point on the boundary of  $Z$  and let  $f_1, \dots, f_\nu, \dots$  be a sequence of points outside  $Z$  approaching  $f$ . We may assume that they lie in  $G(\rho, \mu)$ . A certain substitution  $S_\nu^{-1}$  of  $\{S\}$  will carry  $f_\nu$  into a point  $f'_\nu = S_\nu^{-1} f_\nu$  of  $Z$ . In consequence of our statement about  $G(\rho, \mu)$ , a certain  $S$  recurs infinitely often among the  $S_\nu$ . This  $S$  is no  $J$ . The points  $f'_\nu$  of the corresponding selected sequence all lie in  $Z_S$ , and so does their limit  $f$ . Consequently any point  $f$  on the boundary of  $Z$  belongs to the common boundary of  $Z$  and an equivalent  $Z_S$ . The domain  $G(\rho, \mu)$  increases when  $\rho$  and  $\mu$  increase, and with  $\rho \rightarrow \infty, \mu \rightarrow \infty$  exhausts the whole  $G$ . At every stage  $G(\rho, \mu)$  has points in common with only a finite number of cells  $Z_S$ . An elementary argument like the one applied before shows that among the common boundaries  $Z|Z_S$  (which exist in finite number only) none but the "walls", the  $(N-1)$ -dimensional ones, matter. The substitutions  $S$  which carry  $Z$  into cells  $Z_S$  bordering on  $Z$  along a wall, together with the  $J$ , generate the entire group  $\{S\}$ .

The Generalized Theorem V proves that all our statements about  $G(\rho, \mu)$  remain true if we define it by the following *linear* inequalities:

$$f(x_1, \dots, x_n) \geq \frac{1}{\rho} g_{kk}$$

whenever the  $x_i$  are integers and  $x_k, \dots, x_n$  have no common factor,

$$f(x_1, \dots, x_n) \geq g_{kk} - \mu g_{11}$$

whenever the  $x_i$  are integers and  $(x_k, x_{k+1}, \dots, x_n) = (1, 0, \dots, 0)$ . Incidentally the estimates may be improved considerably if we carry out the comparison of the coefficients  $g_{kk}, g'_{kk}$  of two equivalent forms  $f, f'$  in  $Z$  and  $G(\rho, \mu)$  respectively for any convex bodies rather than ellipsoids by the procedure of § 6.

Throughout we have studied  $Z$  relative to  $G$ ; this is really the more natural standpoint. However, the results I and I\*\* are true even relative to the space  $R$  of all quadratic forms.  $Z$  is not closed in  $R$ ; there may be boundary points  $f$  of  $Z$  relative to  $R$  which are not positive forms. I maintain that for such a form the equation  $g_{11} = 0$  holds. Since  $f$  is non-negative, it can be brought into the form

$$\xi_1^2 + \dots + \xi_m^2 \quad (m < n)$$

by a transformation with real coefficients. We can determine a lattice point  $x \neq 0$  for which the  $m$  linear forms  $\xi_1, \dots, \xi_m$  are in absolute value less than any preassigned positive number  $\epsilon$ , either by Minkowski's theorem (1.1) for parallelotopes, or by a simple application of Dirichlet's principle of the distribution of  $\nu+1$  objects over  $\nu$  boxes. Hence  $g_{11} > 0$  would contradict the inequality  $f(x) \geq g_{11}$  which holds for all lattice points  $x \neq 0$ , and  $Z$  relative to  $R$  is defined by the inequalities described under I\* and I\*\* together with  $g_{11} \geq 0$ . But even the addition of  $g_{11} \geq 0$  becomes superfluous if  $n \geq 2$ . The substitution

$$x_1 = 1, \quad x_2 = \pm 1, \quad x_3 = \dots = x_n = 0$$

yields the inequality

$$g_{11} \pm 2g_{12} + g_{22} \geq g_{22}$$

for any reduced form, or  $2|g_{12}| \leq g_{11}$ . The same holds for  $g_{13}, \dots, g_{1n}$ . Hence the boundary form  $f$  even satisfies the  $n$  equations

$$(8.8) \quad g_{11} = g_{12} = \dots = g_{1n} = 0.$$

The fact that  $Z$  borders on the plane  $g_{11} = 0$  only along this linear variety (8.8) of  $n$  dimensions less, proves our point.

Minkowski's investigations set out from the reduction of quadratic forms. Owing to some "unexpected difficulties" the second half of his book *Geometrie der Zahlen* never appeared, and fourteen years after the discovery of the basic idea of the geometry of numbers he published his theory of reduction in a strictly arithmetical form, without a hint of geometric approach. What was the snag he struck? After having taken the main hurdle with the general inequality (1.2), did he fail to notice the simple remark that led us on to (6.2), or was he unable to unravel the

inductive kink by which in his final publication the establishment of the inequality (6.2) is tied to the selection of a finite number of inequalities 3? It is now idle to speculate. But of him it might be said as of Saul that he went out to look after his father's asses and found a kingdom.

For quadratic forms there is a simple algebraic method of deriving (1.2) from (1.1). Minkowski describes it in *Geometrie der Zahlen*, § 51; and very likely this suggested to him the general inequality (1.2). Hence every improvement of the constant  $\gamma_n$  in the inequality for positive forms

$$M^n \leq \gamma_n D$$

in which  $M$  is the minimum of  $f(x)$  for all lattice points  $x \neq 0$  entails a proportional improvement in (1.2) and for the constants  $\mu_n$  and  $\rho_n$  in the estimates (6.2), (7.1) for ellipsoids. Our value was  $\gamma_n = (2^n/\omega_n)^2$ . Blichfeldt† succeeded in improving it by the factor

$$(1 + \frac{1}{2}n)^2 \cdot 2^{-n}.$$

This leads to the following  $\lambda_n$ :

$$(8.9) \quad 1/\lambda_n = 2^n (1 + \frac{1}{2}n)^2 \omega_n^{-2} \cdot (\frac{9}{4})^{\frac{1}{4}(n-1)(n-2)}.$$

Remak recently obtained a better estimate‡ which differs from (8.9) essentially by having  $\frac{5}{4}$  instead of  $\frac{9}{4}$ . However, the chief interest of our constant  $\mu_n$  in (6.2) lies in its validity for all convex solids whatsoever.

Theorem VI together with Blichfeldt's improved  $\gamma_n$  shows that every positive form has an equivalent one  $f$  satisfying the relation

$$\lambda_n^* g_{11} \cdots g_{nn} \leq D,$$

where 
$$1/\lambda_n^* = (1 + \frac{1}{2}n)^2 \cdot 2^{-n} \omega_n^{-2} \rho_n^2.$$

Such an estimate is best suited to prove that there is only a finite number of classes of equivalent positive forms with integral coefficients.

The Institute for Advanced Study,  
Princeton, N.J., U.S.A.

† *Trans. American Math. Soc.*, 15 (1914), 227–235; cf. R. Remak, *Math. Zeitschrift*, 26 (1927), 694–699, and Blichfeldt, *Math. Annalen*, 101 (1929), 605–608.

‡ *Compositio Math.*, 5 (1938), 368–391.